

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 17:14:36 ; Search time 74.2239 Seconds
(without alignments)
504.414 Million cell updates/sec

Title: US-09-807-933B-11

Perfect score: 1895

Sequence: 1 MKFSIIIASALLLAASSTYAA.....TFKAVTCPAEIIIAKTGCERK 346

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata/2/pubpaa/US08_NEW PUB.pdb.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW PUB.pdb.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW PUB.pdb.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pdb.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW PUB.pdb.*
- 6: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pdb.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pdb.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pdb.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_NEW PUB.pdb.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pdb.*
- 11: /cgn2_6/prodata/2/pubpaa/US10_NEW PUB.pdb.*
- 12: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pdb.*
- 13: /cgn2_6/prodata/2/pubpaa/US60_NEW PUB.pdb.*
- 14: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pdb.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	696.5	36.8	201	9	US-09-261-329-4
2	686.5	36.2	299	9	US-10-007-521-12
3	683.5	36.1	203	9	US-09-261-329-9
4	683.5	36.1	222	9	US-10-007-521-14
5	683.5	36.1	294	9	US-10-007-521-24
6	680	35.9	349	9	US-10-007-521-10
7	679.5	35.9	201	9	US-09-261-329-5
8	678.5	35.8	225	9	US-10-007-521-2
9	678.5	35.8	297	9	US-10-007-521-4
10	678.5	35.8	308	9	US-10-007-521-6
11	672.5	35.5	205	9	US-09-261-329-7
12	668	35.3	310	9	US-10-007-521-22
13	664	34.4	202	9	US-09-261-329-3
14	652	34.4	305	10	US-09-735-787-2
15	651.5	34.4	235	1	US-08-841-636A-31
16	651	34.0	202	9	US-09-261-329-1
17	644.5	34.0	376	10	US-09-735-787-4
18	640.5	33.8	203	9	US-09-261-329-6
19	634.5	33.5	226	9	US-10-007-521-16

20	634.5	33.5	293	9	US-10-007-521-20	Sequence 20, Appl
21	634.5	33.5	298	9	US-10-007-521-18	Sequence 18, Appl
22	629.5	33.2	203	9	US-09-261-329-8	Sequence 8, Appl
23	620	32.7	295	9	US-10-007-521-8	Sequence 8, Appl
24	616	32.5	202	9	US-09-261-329-2	Sequence 2, Appl
25	468.5	24.7	235	9	US-09-261-329-10	Sequence 10, Appl
26	456.5	24.1	211	9	US-09-261-329-11	Sequence 11, Appl
27	388	20.5	138	9	US-10-007-521-26	Sequence 26, Appl
28	2750	12.1	18636	9	US-10-123-155-85	Sequence 85, Appl
29	228.5	12.1	18636	9	US-10-073-913-17	Sequence 17, Appl
30	227	12.0	2316	9	US-10-123-155-69	Sequence 69, Appl
31	227	12.0	3060	9	US-10-184-644-337	Sequence 337, Appl
32	227	12.0	3060	9	US-10-184-634-337	Sequence 337, Appl
33	226.5	12.0	493	9	US-10-197-294A-2	Sequence 2, Appl
34	223	11.8	4060	9	US-10-123-155-197	Sequence 197, Appl
35	222.5	11.7	4440	9	US-10-174-590-525	Sequence 525, Appl
36	222.5	11.7	4440	9	US-10-176-758-525	Sequence 525, Appl
37	222.5	11.7	4440	9	US-10-175-737-525	Sequence 525, Appl
38	222.5	11.7	4440	9	US-10-173-706-525	Sequence 525, Appl
39	222.5	11.7	4440	9	US-10-175-738-525	Sequence 525, Appl
40	222.5	11.7	4440	9	US-10-175-752-525	Sequence 525, Appl
41	222.5	11.7	4440	9	US-10-176-482-525	Sequence 525, Appl
42	222.5	11.7	4440	9	US-10-176-757-525	Sequence 525, Appl
43	222.5	11.7	4440	9	US-10-176-913-525	Sequence 525, Appl
44	222.5	11.7	4440	9	US-10-180-552-525	Sequence 525, Appl
45	222.5	11.7	4440	9	US-10-180-557-525	Sequence 525, Appl

ALIGNMENTS

RESULT 1

US-09-261-329-4
; Sequence 4, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schuelein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ IDS: 26
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 201
; TYPE: PRT
; ORGANISM: Cellulase variants
US-09-261-329-4

Query Match 36.8%; Score 696.5; DB 9; Length 201;

Best Local Similarity 59.6%; Pred. No. 1.5e-47;

Matches 121; Conservative 28; Mismatches 51; Indels 3; Gaps 1;

QY	143	GNRITRYWDCCKPCSCAMDGKASVTKPVLTCAKDGVSRIGSDVQSCGCGQAYMCDNOP	202
Db	1	GSCKSTRYWDCCCKPCSCAMSGKASVNRPLVACDANNPLNDANVKSCKDCGSAITCANNSP	60
QY	203	WVYNDLAVGFAAASIGSAGAFCCGCVLFTNTAVAGKFFVQVTTNGDLSNHF	262
Db	61	WAVNDLAVGFAAATKLSGGTSSWCACALYITFGSPVSGKTLVQVSTGDLGNSHFD	120
QY	263	LQMPGGGVYFNGCOSOMTNTDGMGARYGGISSTSECCKLPTOLQAGCKWFGWFKNAD	322
Db	121	LNPFGGVSLFGCKREFGGLP---GAQYGGISSRSECDSFPAALKPGCWRFDFWKNAD	177
QY	323	NPEVTFKAVTCPAEIIIAKTGCER	345

Wed Jun 18 17:55:03 2003

Db 178. NPEFTFKOVQCPSELTSRTGCKR 200

RESULT 2

US-10-007-521-12

Sequence 12, Application US/10007521

Publication No. US20030054539A1

GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takegi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESSES:

ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US20030054539A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/007,521

FILING DATE: 10-Dec-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/651,136

FILING DATE: 21-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 4366.200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 299 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLSCULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-10-007-521-12

Query Match 36.2%; Score 686.5; DB 9; Length 299;

Best Local Similarity 53.6%; Pred. No. 1.4e-46;

Matches 125; Conservative 31; Mismatches 66; Indels 11; Gaps 2;

QY 113 KTTTPTTKTTAAATTTSSNTGYSPISGGFSGNGRTRTYWDCCKPSCAWDGKASVTKPVLT 172

Db 2 RSTPVRTTLAALPL-----VASAASGSGSQSTRYWDCKPSCAWPGKAASQPVYA 53

QY 173 CAKDGVSRLGSDVQSGCVGGQAYMNDNOPWVNDLDTLAYGFAAASLGSAGSAFCCGVE 232

Db 54 CDANFQRLSDFNVSQGCNGSAYSCADQTFWAVNDNLAYGFAATSIAGSSSSWCCACYA 113

QY 233 LTFNTAVAGKFFVQVNTTGDLSNTHFDLQMPGGGVYFNGCQSQMNTNTDNGARYG 292

Db 114 LTFSGVAGKTNVWVQSTGGDLGNSQFDIAMPGGGVIFNGCQQFGLP---GAQVG 170

QY 293 GISSISBCDKLPTQLQAGCKWRFKFNADNPVTFKAVTCAPIIAKTGCR 345

Db 171 GISSRQDCDSFPAPLPGCQWRDFQNDNPFTFTFQQVQCQPAEIVARSGCKR 223

Db 178. NPEFTFKOVQCPSELTSRTGCKR 200

RESULT 3

US-09-261-329-9

Sequence 9, Application US/09261329

Publication No. US20030092097A1

GENERAL INFORMATION:

APPLICANT: Andersen, Kim

APPLICANT: Schulein, Martin

APPLICANT: Christiansen, Lars

APPLICANT: Damgaard, Bo

APPLICANT: Von Der Osten, Claus

TITLE OF INVENTION: Cellulase Variants

FILE REFERENCE: 4887.204-US

CURRENT APPLICATION NUMBER: US/09/261,329

EARLIER FILING DATE: 1999-03-03

EARLIER APPLICATION NUMBER: 1013/96

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 9

LENGTH: 203

TYPE: PRT

ORGANISM: Cellulase variants

US-09-261-329-9

Query Match 36.1%; Score 683.5; DB 9; Length 203;

Best Local Similarity 58.6%; Pred. No. 1.6e-46;

Matches 119; Conservative 30; Mismatches 51; Indels 3; Gaps 2;

QY 144 NGRTRYWDCCKPSCAWDGKASVTKPVLTCAKDGVSRLGSD-VQSGCVGGQAYMNDNOP 202

Db 2 SGVTRYWDCCKPSCAWTGKASVSKFVGTCINDNDAQTPTDLLKSSDCGSAAYCSNQGP 61

QY 203 WYVNDLDTLAYGFAAASLGSAGSAFCCGVELTFTNTAVAGKFFVQVNTTGDLSNTHFD 262

Db 62 WAVNDSLSYGFAAAKLSGKQETDWCQCYKLTFTSTAVSGKQMVQITNTGGDLGNHFD 121

QY 263 LQMPGGGVYFNGCQSQMNTNTDNGARYGGISSISECDKLPQLQAGCKWRFKFNAD 322

Db 122 IAMPGGGVIFNGCSQKW--NGINLGNQIGFTDRSQCATLPSKWQASCNWRFDFENAD 179

QY 323 NPEVTFKAVTCAPIIAKTGCR 345

Db 180 NPTVDWEPTVCPQELVARTGCSR 202

RESULT 4

US-10-007-521-14

Sequence 14, Application US/10007521

Publication No. US20030054539A1

GENERAL INFORMATION:

APPLICANT: Schulein, Martin

Andersen, Lene N.

Lassen, Soren F.

Kauppinen, Markus S.

Lange, Lene

Nielsen, Ruby I.

Ihara, Michiko

Takegi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases

NUMBER OF SEQUENCES: 109

CORRESPONDENCE ADDRESSES:

ADDRESSEE: No. US20030054539A1o No. US20030054539A1disk of No. US20030054539A1

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 349 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-007-521-10

Query Match 35.9%; Score 680; DB 9; Length 349;
Best Local Similarity 57.4%; Pred. No. 5.6e-46;
Matches 117; Conservative 32; Mismatches 53; Indels 2; Gaps 1;

Qy 142 SGNRTTRYWDCCKPSCAWDGKASVTKPVLTCADKGVSRIGSDVQSGCGVQAYMCDNQ 201
Db 22 SGKGHTRYWDCCKTSCAWEGKASVSEBVLTCNKQDNPIVDANARSGCDGGAFACNNNS 81

Qy 202 PWVNDLLAYGFAAASLGASAGAFCCGCGCYELTETNTAVAGKKFVVQVNTGGDLSTNHF 261
Db 82 PWAVSEDLAYGFAATALSGGTGEGSWCCACVAITFTSGFVAGKRWVQSTNITGGDLNNHF 141

Qy 262 DLQMPGGGVYFNGCQSQWNTNTDVGARYGGISSISECDKLPLOLAGCKWRFGWPKNA 321
Db 142 DLMIPGGGLGIFDGCAPGQLLP--GERYGVSSRSQCDGMPPELLKXGQWRFDWFKNS 199

Qy 322 DNPEVTFKAVTCPAEIIAKTGCR 345
Db 200 DNPDIPEFQVQCPKELIAVSGCVR 223

RESULT 7

US-09-261-329-5
Sequence 5, Application US/09261329
Publication No. US20030092097A1
GENERAL INFORMATION:
APPLICANT: Andersen, Kim
APPLICANT: Schulein, Martin
APPLICANT: Christiansen, Lars
APPLICANT: Damgaard, Bo
APPLICANT: Von Der Osten, Claus
TITLE OF INVENTION: Cellulase Variants
FILE REFERENCE: 4887.204-US
CURRENT APPLICATION NUMBER: US/09/261,329
CURRENT FILING DATE: 1999-03-03
EARLIER APPLICATION NUMBER: 1013/96
EARLIER FILING DATE: 1996-09-17
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 201
TYPE: PRT
ORGANISM: Cellulase variants
US-09-261-329-5

Query Match 35.9%; Score 679.5; DB 9; Length 201;
Best Local Similarity 58.6%; Pred. No. 3.3e-46;
Matches 119; Conservative 27; Mismatches 54; Indels 3; Gaps 1;

Qy 143 GNRTRYWDCCKPSCAWDGKASVTKPVLTCADKGVSRIGSDVQSGCGVQAYMCDNQ 202
Db 1 GSGQSTRYWDCKPSCAWPGKAASQPVVYACDANFRLSDNFVQSGCNGGSAISCADQTP 60

Qy 203 WVVNDLLAYGFAAASLGASAGAFCCGCGCYELTETNTAVAGKKFVVQVNTGGDLSTNHF 262

Db 61 WAVNDNLAYGFAAATSIAGGSESSMCCACALTFTTSGPVAGKTMVQVSTTGGDLGNOFD 120

Qy 263 LQMPGGGVYFNGCQSQWNTNTDVGARYGGISSISECDKLPLOLAGCKWRFGWPKNA 322

Db 121 IAMPGGGVYFNGCQSQGGLP--GAOYGGISSRDQCDSPAPLKPCCQWRFDWFOAD 177

Qy 323 NPEVTFKAVTCPAEIIAKTGCR 345

Db 178 NPTFTFQVQCPAEIVARSCKR 200

RESULT 8

US-10-007-521-2
Sequence 2, Application US/10007521
Publication No. US20030054539A1
GENERAL INFORMATION:
APPLICANT: Schulein, Martin
Andersen, Lene N.
Lassen, Soren F.
Kauppinen, Markus S.
Lange, Lene
Nielsen, Ruby I.
Ihara, Michiko
Takagi, Shinobu
TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESSES: No. US20030054539A1o No. US20030054539A1disk of No. US20030054
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-007-521-2

Query Match 35.8%; Score 678.5; DB 9; Length 225;
Best Local Similarity 54.3%; Pred. No. 4.5e-46;
Matches 121; Conservative 26; Mismatches 69; Indels 7; Gaps 4;

Qy 129 SSNTGYSPIS---GGFSGNGRTRYWDCCKPSCAWDGKASVTKPVLTCADKGVSRIGS 183

Db 4 SATTGTFLALPVLALDQLSGIGIGQITRYWDCCKPSCAWPGKGP--SSPVQACDKNDNPLNDGG 62

Qy 184 DVQGGC-VGGQAYMCDNQWVNDLLAYGFAAASLGASAGAFCCGCGCYELTETNTAVAG 242

Db 63 STRSGCDAGGSAYMCCSQSPWAVSDELSYGMVAVKLAGSSESQWCCACCYELTFTTSGPVAG 122

[illegible]

```

RESULT 10
US-10-007-521-6
; Sequence 6, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; Andersen, Lene N.
; Lassen, Soren F.
; Kauppinen, Markus S.
; Lange, Lene
; Nielsen, Ruby I.
; Ihara, Michiko
; Takagi, Shinobu
; TITLE OF INVENTION: No. US20030054539A1 Endoglucanases
; NUMBER OF SEQUENCES: 109
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. US20030054539A1 No. US20030054539A1
; STREET: 405 Lexington Avenue, 64th Floor
;

```

COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-007-521-6

[illegible]

QY 303 LPTQLQAGCKWRFGWFKVADNPEVTFKAVTCTPAEIIAKTCER 345
Db 183 FPEALPGCNWRFDFWQADNPSVTFQEVACPSELTSKSGCSR 225

RESULT 11

US-09-261-329-7
; Sequence 7, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schulein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Cellulase variants
US-09-261-329-7

Query Match 35.5%; Score 672.5; DB 9; Length 205;
Best Local Similarity 57.1%; Pred. No. 1.2e-45;
Matches 117; Conservative 23; Mismatches 62; Indels 3; Gaps 3;

QY 143 GNCETTRYWDCCKPSCAWDGKASVTKPVLTCAK-DGVSRLGSDVQSGC-VGGOAYMCDN 200
Db 1 GIGQTTTRYWDCCKPSCAWDPGKP-SSPVQACDKNDNFNDGGSTRSCDAGGAYMCSQ 59
QY 201 QPWWNDLALYGAFAAASLGASAGAFCCGCVYELTFTNTAVAGKFFVQVVTNTGDDLSTNH 260
Db 60 SPWAVSDELSYGMAAVKLAGSESQWCCACVYELTFTSGPVAGKMWIQAINTGGDLGDNH 119
QY 261 FDLQMPGGGVYFNGCOSQWNTNDGWGARYGGISSISECDKLPQLQAGCKWRFGWFKN 320
Db 120 FDLAIPGGGGVIFNACTDQYGAPPNGWGDYGGIHSKECESFPPEALKPGCNWRFDFWQ 179
QY 321 ADNPEVTFKAVTCTPAEIIAKTCER 345
Db 180 ADNPSVTFQEVACPSELTSKSGCSR 204

RESULT 12

US-10-007-521-22
; Sequence 22, Application US/10007521
; Publication No. US20030054539A1
; GENERAL INFORMATION:
; APPLICANT: Schulein, Martin
; Andersen, Lene N.
; Lassen, Soren F.
; Kauppinen, Markus S.
; Lange, Lene
; Nielsen, Ruby I.
; Ihara, Michiko
; Takagi, Shinobu

TITLE OF INVENTION: No. US20030054539A1el Endoglucanases
NUMBER OF SEQUENCES: 109
CORRESPONDENCE ADDRESS:
ADDRESS: No. US20030054539A1o No. US20030054539A1disk of No. US20030054539A1
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/007,521
FILING DATE: 10-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/651,136
FILING DATE: 21-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4366.200-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 310 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-10-007-521-22

Query Match 35.3%; Score 668; DB 9; Length 310;

Best Local Similarity 57.1%; Pred. No. 4.3e-45;
Matches 117; Conservative 29; Mismatches 55; Indels 4; Gaps 3;

QY 142 SGNGRITRYWDCCKPSCAWDGKASVTKPVLTCAGDVSRLGSDVQSGC-VGGOAYMCDN 200
Db 19 SGTGRITRYWDCCKPSCGWDKASVSQPVKTCDRNN-NPLASTARSGCDNSGVAYTCNDN 77
QY 201 QPWWNDLALYGAFAAASLGASAGAFCCGCVYELTFTNTAVAGKFFVQVVTNTGDDLSTNH 260
Db 78 QPWWNDLALYGAFAAATAFSGGSEASWCCACVYELTFTSGPVAGKMWVQVSTNTGDLJSGNH 137
QY 261 FDLQMPGGGVYFNGCOSQWNTNDGWGARYGGISSISECDKLPQLQAGCKWRFGWFKN 320
Db 138 FDLMPGGGLGIFDGTTPQWGVSF--GNRYGGTTSRSCSQSPALOPCNCWRYDFWFD 195
QY 321 ADNPEVTFKAVTCTPAEIIAKTCER 345
Db 196 ADNPDVSWRRVQCPAALTDRTGCR 220

RESULT 13

US-09-261-329-3
; Sequence 3, Application US/09261329
; Publication No. US20030092097A1
; GENERAL INFORMATION:
; APPLICANT: Andersen, Kim
; APPLICANT: Schulein, Martin
; APPLICANT: Christiansen, Lars
; APPLICANT: Damgaard, Bo
; APPLICANT: Von Der Osten, Claus
; TITLE OF INVENTION: Cellulase Variants
; FILE REFERENCE: 4887.204-US
; CURRENT APPLICATION NUMBER: US/09/261,329
; CURRENT FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 1013/96
; EARLIER FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 202
; TYPE: PRT
; ORGANISM: cellulase variants
US-09-261-329-3

Query Match 34.4%; Score 652; DB 10; Length 305;
Best Local Similarity 56.6%; Pred. No. 7.7e-44;
Matches 116; Conservative 25; Mismatches 60; Indels 4; Gaps 2;
REGISTRATION NUMBER: 417306
REFERENCE/DOCKET NUMBER: 1716.0510005/MAC/TJS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540

```
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 235 amino acids
;   TYPE: amino acid
;   STRANDEDNESS:
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   ORIGINAL SOURCE:
;   ORGANISM: Melanocarpus albomyces
;   STRAIN: ALK04237
;   FEATURE:
;   NAME/KEY: Protein
;   LOCATION: 1..235
;   OTHER INFORMATION: /label= 20K-cellulase
;
US-08-841-636A-31

Query Match      34.4%; Score 651.5; DB 1; Length 235;
Best Local Similarity 54.6%; Pred. No. 6.3e-44;
Matches 113; Conservative 27; Mismatches 64; Indels 3; Gaps 1;

Qy 139 GGFSGNGRTTRYWDCKPSCAWDGKASVTKPVLTCAKDGVSRGLSDVQSGCVGQAYMCN 198
Db 18 GALAANGQSTRYWDCKPSCGWRGKGPVNPQPVYSCDANFORIHDFAVSGCEGPAFSCA 77
Qy 199 DNQPVVYNDLALYGAFAASLGSAGASAFCCGCGYELTFTNTAVAGKFEVQVNTNGDDLST 258
Db 78 DHSPWAINDLNLSYGFAATALSQGQTEESWCCACVALTFTSGFVAGKTMVQVSTSTGGDLGS 137
Qy 259 NHFDLQMPGGGVYFNGCQSQWNTNTDQWARYGGISSISECDKLPLOAGCKWRFGWF 318
Db 138 NHFDLNIPIGGGVGLFDGCTPQFGGLP---GARYGGISSRQECDSFPPEPLKPGCOWRFDWF 194
Qy 319 KNADNPEVTPKAVTCPAEIIAKTGCR 345
Db 195 QNADNPSFTFERVQCPPELVARTGCR 221
```

Search completed: June 18, 2003, 17:44:45
Job time : 75.2239 secs